J Arid Land (2025) 17(1): 93–111
doi: 10.1007/s40333-024-0066-3; CSTR: 32276.14.JAL.02400663

Science Press Springer-Verlag

# An improved GCN–TCN–AR model for $PM_{2.5}$ predictions in the arid areas of Xinjiang, China

CHEN Wenqian[1*], BAI Xuesong[1], ZHANG Na[1], CAO Xiaoyi[2]

[1] School of Information and Control Engineering, Qingdao University of Technology, Qingdao 266520, China;
[2] Key Laboratory for Semi-Arid Climate Change of the Ministry of Education, College of Atmospheric Sciences, Lanzhou University, Lanzhou 730000, China

**Abstract:** As one of the main characteristics of atmospheric pollutants, $PM_{2.5}$ severely affects human health and has received widespread attention in recent years. How to predict the variations of $PM_{2.5}$ concentrations with high accuracy is an important topic. The $PM_{2.5}$ monitoring stations in Xinjiang Uygur Autonomous Region, China, are unevenly distributed, which makes it challenging to conduct comprehensive analyses and predictions. Therefore, this study primarily addresses the limitations mentioned above and the poor generalization ability of $PM_{2.5}$ concentration prediction models across different monitoring stations. We chose the northern slope of the Tianshan Mountains as the study area and took the January–December in 2019 as the research period. On the basis of data from 21 $PM_{2.5}$ monitoring stations as well as meteorological data (temperature, instantaneous wind speed, and pressure), we developed an improved model, namely GCN–TCN–AR (where GCN is the graph convolution network, TCN is the temporal convolutional network, and AR is the autoregression), for predicting $PM_{2.5}$ concentrations on the northern slope of the Tianshan Mountains. The GCN–TCN–AR model is composed of an improved GCN model, a TCN model, and an AR model. The results revealed that the $R^2$ values predicted by the GCN–TCN–AR model at the four monitoring stations (Urumqi, Wujiaqu, Shihezi, and Changji) were 0.93, 0.91, 0.93, and 0.92, respectively, and the RMSE (root mean square error) values were 6.85, 7.52, 7.01, and 7.28 μg/m³, respectively. The performance of the GCN–TCN–AR model was also compared with the currently neural network models, including the GCN–TCN, GCN, TCN, Support Vector Regression (SVR), and AR. The GCN–TCN–AR outperformed the other current neural network models, with high prediction accuracy and good stability, making it especially suitable for the predictions of $PM_{2.5}$ concentrations. This study revealed the significant spatiotemporal variations of $PM_{2.5}$ concentrations. First, the $PM_{2.5}$ concentrations exhibited clear seasonal fluctuations, with higher levels typically observed in winter and differences presented between months. Second, the spatial distribution analysis revealed that cities such as Urumqi and Wujiaqu have high $PM_{2.5}$ concentrations, with a noticeable geographical clustering of pollutions. Understanding the variations in $PM_{2.5}$ concentrations is highly important for the sustainable development of ecological environment in arid areas.

**Keywords:** air pollution; $PM_{2.5}$ concentrations; graph convolution network (GCN) model; temporal convolutional network (TCN) model; autoregression (AR) model; northern slope of the Tianshan Mountains

**Citation:** CHEN Wenqian, BAI Xuesong, ZHANG Na, CAO Xiaoyi. 2025. An improved GCN–TCN–AR model for $PM_{2.5}$ predictions in the arid areas of Xinjiang, China. Journal of Arid Land, 17(1): 93–111. https://doi.org/10.1007/s40333-024-0066-3; https://cstr.cn/32276.14.JAL.02400663

*Corresponding author: CHEN Wenqian (E-mail: chimmyqu@yeah.net)
Received 2024-09-18; revised 2024-11-21; accepted 2024-11-25

## 1 Introduction

With economic development, air pollution has become an environmental problem that urgently needs to be solved in China (Zhang et al., 2024). Owing to the serious impact of air pollution on human health and the living environment, it has been a concern and has become the focus of common attention in all sectors of society (Ma et al., 2020). $PM_{2.5}$ is a typical atmospheric pollutant that refers to the fine particles in the air with the diameter of less than or equal to 2.5 μm. Various toxic substances are easily attached to $PM_{2.5}$, gradually accumulate at the end of the respiratory tract, and affect the other parts of the body through air exchange in the lungs, which has a major impact on human health (Sivarethinamohan et al., 2021). Although the number of $PM_{2.5}$ monitoring stations in China is small and unevenly distributed, the concentrations of $PM_{2.5}$ at these stations can not only represent the general trend of $PM_{2.5}$ concentrations in a given area, but also predict the overall change in $PM_{2.5}$ concentrations relative to the area's surface characteristics when combined with local aerosol parameters. Therefore, developing a highly accurate and stable prediction model for the $PM_{2.5}$ concentrations at monitoring stations can help us better understand the current characteristic distributions and future change trends of $PM_{2.5}$ concentrations, thus formulating targeted environmental protection and policy measures (Yu et al., 2020; Wu et al., 2021; Ye et al., 2023).

With the widespread application of machine learning technology, the related models, such as the autoregressive integrated moving average (Bhatt et al., 2021), random forest (Xia et al., 2020), and support vector machine (Zhou et al., 2019), have been used to predict $PM_{2.5}$ concentrations at monitoring stations. In recent years, deep learning technology has developed rapidly. Compared with machine learning, deep learning uses deep neural networks to effectively capture nonlinear relationships in data through hierarchical feature representation (Saha et al., 2021). Researchers have applied deep learning models to prediction tasks (Bai and Shen, 2019; Mohamed, 2019; Dong et al., 2021). For example, Mohamed (2019) used artificial neural network to predict the air quality index of Ahvaz, Iran, and proved its applicability through model comparison experiments. Bai and Shen (2019) investigated the application of long short-term memory (LSTM) neural network in predicting $PM_{2.5}$ concentrations, and revealed that despite its excellent performance in time series prediction, the high memory requirement of LSTM neural network in processing large-scale data limits the breadth of its application. Xing et al. (2022) proposed a model that combines convolutional neural network (CNN) and LSTM neural network, which is excellent in terms of prediction accuracy and generalizability. However, this integrated model still faces the problems of high memory requirements and limited effectiveness in addressing complex spatial features. To address these issues, Lu and Li (2023) proposed a novel multisite prediction model that integrates a CNN model and bidirectional LSTM (BiLSTM) neural network and adjusts the hyperparameters through Bayesian optimization to improve the generalization ability of the model. However, this multisite prediction model still fails to effectively address the uneven distribution of $PM_{2.5}$ monitoring stations (Lu and Li, 2023). In practical applications, the graph convolutional network (GCN) has attracted the attention of researchers because of its ability to extract features from non-Euclidean space (Zhao et al., 2021). The GCN is particularly suitable for capturing the spatial features of $PM_{2.5}$ concentrations. The hybrid neural network model proposed by Mohammadzadeh et al. (2024), which integrates the GCN and LSTM hybrid neural network models, successfully captures the spatiotemporal features of $PM_{2.5}$ concentrations and achieves good prediction results, but it still faces data sparsity and computational complexity problems in practical applications. The TCN model can effectively obtain historical $PM_{2.5}$ concentration data through a causal convolutional system and can use only historical data information for forecasting $PM_{2.5}$ concentrations without considering the $PM_{2.5}$ information in the future (Jiang et al., 2021). However, the methods and studies based on the TCN model are not perfect, and some shortcomings remain. The TCN model currently under study focuses mainly on the modeling of time series, so capturing the spatiotemporal characteristics of $PM_{2.5}$ concentrations remains a challenge.

The northern slope of the Tianshan Mountains has been identified as one of the key development areas in Xinjiang Uygur Autonomous Region, China. However, very few $PM_{2.5}$ monitoring stations exist in this region, which makes it challenging to conduct comprehensive analyses and predictions. The analysis and accurate predictions of $PM_{2.5}$ concentrations can help support the sustainable development of the northern slope of the Tianshan Mountains. On the basis of the current research status, this paper proposes a new spatiotemporal prediction model of $PM_{2.5}$ concentrations, with the goal of achieving precise $PM_{2.5}$ predictions at multiple monitoring stations on the northern slope of the Tianshan Mountains. This spatiotemporal $PM_{2.5}$ prediction model integrates three model structures: an improved GCN model, a TCN model, and an autoregression (AR) model, which is called the GCN–TCN–AR model. The GCN model can address complex spatial features and extract complex spatial features of $PM_{2.5}$ concentrations. The TCN model is a deep learning model for processing time series data that can effectively extract complex relationships and features in $PM_{2.5}$ time series data. When combined with the linear model, the GCN–TCN–AR model can efficiently capture the spatiotemporal characteristics of $PM_{2.5}$ concentrations and effectively improve the accuracy of regional $PM_{2.5}$ predictions. The innovations of this study are as follows: (1) constructing a new spatiotemporal model for $PM_{2.5}$ predictions, which can efficiently extract the spatiotemporal characteristics of $PM_{2.5}$ concentrations and increase the accuracy of the model predictions by combining both linear and nonlinear models; and (2) proposing an improved GCN model. Specifically, we combined the GCN model with the residual model, so the improved GCN model integrates the characteristics of the graph convolution operation and residual connection, and can effectively learn complex relationships and feature representations in data, thus improving the training effect and performance of the model. The GCN–TCN–AR model can not only accurately predict $PM_{2.5}$ concentrations but also reveal the spatial dependencies and temporal evolution patterns between different regions, making it highly valuable for environmental quality monitoring and public health early warning applications.

## 2 Materials and methods

### 2.1 Study area

This study focuses on the northern slope of the Tianshan Mountains (79°53′06″–88°58′19″E, 42°55′19″–46°12′42″N; Fig. 1) located in Xinjiang, China. The terrain of the study area presents high altitudes in the south and low altitudes in the north. The region is characterized by a temperate continental semi-arid climate, with strong sunlight and evaporation, making the atmospheric ecological environment very fragile. The annual precipitation is relatively low, with significant seasonal differences. The annual precipitation in the mountainous areas of this region is 500 mm; the months with the higher precipitation mostly occur from May to July, and the months with the lower precipitation occurs in January.

### 2.2 Model fundamentals

#### 2.2.1 GCN model

The GCN model is an effective method for dealing with irregular geometry problems and improving the ability to analyze complex spatial relationships (Wang et al., 2024b). However, since $PM_{2.5}$ concentration presents spatial distribution characteristics in reality, the GCN model utilizes local connectivity and internode relationships within the graph for efficient spatial feature extraction to improve its prediction performance (Gao and Li, 2021). This makes it ideal for extracting the spatial distribution characteristics of $PM_{2.5}$ concentrations. The construction of the graph is crucial when the GCN model is used to predict $PM_{2.5}$ concentrations at monitoring stations. The construction of the graph considers the spatial relationships between monitoring stations and connectivity information to help the model capture the spatial features between monitoring stations. Combined with the definition of the graph, the spatial relationship between

different monitoring stations can be expressed as an undirected graph ($G$): $G=(N, E, A)$, where $N$ is the number of monitoring stations, $(v_i, v_j)\in E$ represents every edge (where $v$ is the monitoring station; and $i$, $j$=1, 2,$\cdots$, $n$), and $A\in R^{N\times N}$ is the adjacency matrix for measuring the spatial correlation between $v_i$ and $v_j$ (where $A$ is the adjacency matrix and $R$ is the spatial correlation between monitoring stations) (Mohammadzadeh et al., 2024). However, because the GCN model focuses on local spatial connectivity, its effectiveness in capturing long-range dependencies (which are crucial for certain datasets) is limited. This shortcoming highlights the need for a model that can integrate both local and global spatial features.

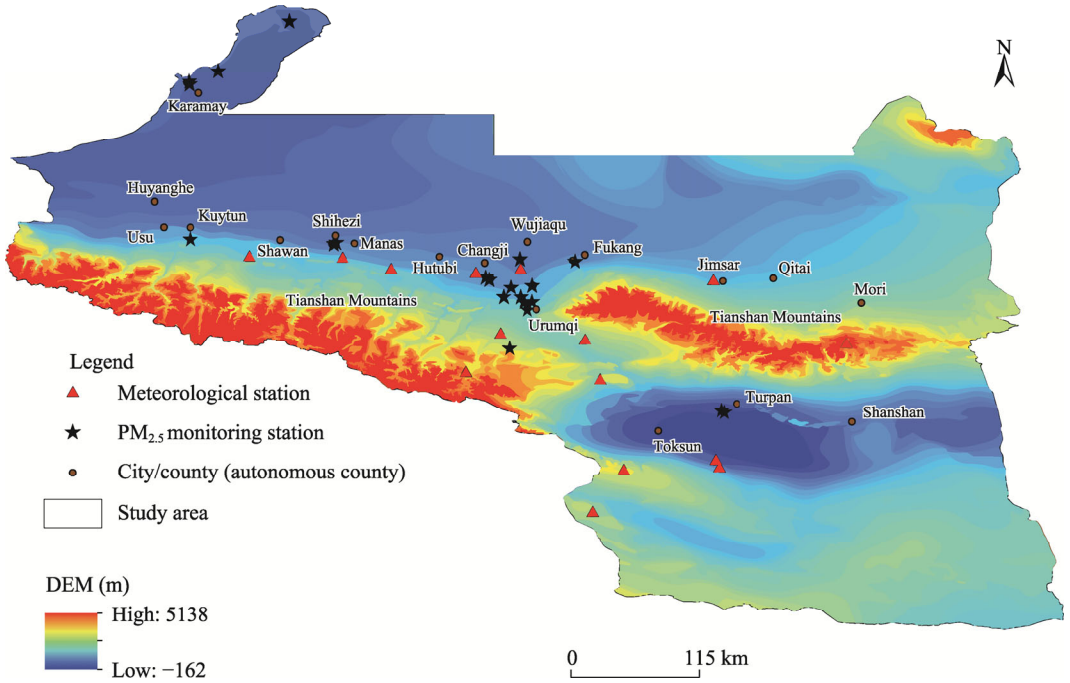


**Fig. 1** Overview of the study area (northern slope of the Tianshan Mountains) based on the digital elevation model (DEM) and the spatial distributions of $PM_{2.5}$ monitoring stations and meteorological stations

In the predictions of $PM_{2.5}$ concentrations, air quality data from different areas can be constructed into a graph structure due to the obvious spatial propagation pattern of air pollutants. The graph convolution residual model can effectively utilize this graph structure to better capture the air pollution propagation relationships between areas (Yao et al., 2023). First, the graph convolution layer is similar to the operation of a CNN model on image data, which can propagate information and feature extraction in image data. This operation captures the relationships and local structure between nodes and generates a new feature representation of the node. Then, the residual connection layer realizes jump connection and information transfer by performing an addition operation between the input features and the features that have been processed by the graph convolution layer (Yu et al., 2020). Such a design can not only help the model pass the gradient better during the training process, avoiding the gradient vanishing or gradient explosion problem, but also make the model easy to train and achieve convergence. By stacking multiple graph convolutional residual models, a deep graph neural network model can be constructed to predict the $PM_{2.5}$ concentrations at monitoring stations using the GCN model.

During the process of $PM_{2.5}$ predictions, the graph convolution residual equations were used to express the relationship between the inputs and outputs of the graph convolution layer instead of learning the outputs directly. In general, for a graph convolution layer with residual connections, the residual formula can be expressed as follows:

$$H^{(r+1)} = f(H^{(r)} + /m(X)(H^{(r)}, Y)), \quad (1)$$

where $H^{(r+1)}$ is the output feature representation of layer $r$+1 and is used as the input to the next layer; $f$ is the activation function applied to the result of the addition operation; $H^{(r)}$ is the input hidden feature representation of layer $r$; $Y$ is the graph structure, which is represented in the form of an adjacency matrix; and $/m(X)(H^{(r)}, Y)$ represents the graph convolution operation, which is a function that combines $H^{(r)}$ with $Y$ to compute the output features. This design allows the network to gradually adjust the input features by learning the residuals to better fit the target.

**2.2.2** TCN model

The TCN model is composed of three core models: causal convolution, extended convolution, and residual connection. It effectively avoids the gradient vanishing or gradient explosion problem encountered in recurrent neural networks, and has the advantages of parallel computation, low memory consumption, and the ability to capture both long-term and short-term temporal features (Zeng et al., 2023). The TCN model is an effective method for extracting the time features of $PM_{2.5}$ concentrations. By stacking multiple convolution layers, the TCN model captures the features at different scales, thereby effectively capturing the long-term complex time dependence. Moreover, the convolutional operations in the TCN model can parallelize the processing of different time steps to achieve faster training and inference. Therefore, the use of a TCN model to extract the temporal features of $PM_{2.5}$ concentrations is crucial for accurately predicting the changes in $PM_{2.5}$ concentrations. While the TCN model effectively captures the long-term complex time dependence, its reliance on fixed convolutional structures may reduce its adaptability to varying temporal patterns, leading to potential inaccuracy in more complex time series data.

The task of modeling sequences involving sequences with long-term complex time dependency becomes challenging because of the linear correlation between the length of historical data captured by causal convolution and the depth of the network. The extended convolution is used in the TCN model to extend the acceptance domain of the network.

The process of calculating the causal convolution (taking $k$=1, 2, 3, 4 and $I$=2 as an example, where $k$ is the time step at a given time and $I$ is the filter size) is shown in Equation 2:

$$F(a) = \sum_{J=0}^{I-1} g(J) \times X_{a-k_I}, \quad (2)$$

where $F(a)$ is a function, in which $a$ is the input sequence; $J$ denotes the iteration variable; $g(J)$ denotes the value of the filter at position $J$; $X$ represents the series; and $a–k_I$ represents the direction of the past.

The time convolution second-order residual model is a deep learning model structure for processing time series data, which combines the properties of the time convolution operation and residual linkage to effectively learn complex relationships and feature representations of time series data (Chen et al., 2018). In the second-order residual connection, the feature extraction and information propagation of time series data are first carried out through a time convolution operation. Temporal convolution operations capture the relationships and patterns between different time points in time series data to generate new representations of temporal features. The second-order residual connection performs the summation operation for the features obtained after the temporal convolution operation and the input features, which realizes jump connection and information transfer (Zhang et al., 2022). This design can help the model better learn the long-term complex time dependency in the time series data and improve the generalizability and prediction accuracy of the model (Ren et al., 2023). By stacking multiple time convolution second-order residual model, a deep time series prediction model can be constructed for accurate prediction and analysis of time series data. The second-order residual connection is more prominent in extracting complex temporal features.

**2.2.3** AR model

The AR model is a commonly used time series prediction method (Zhen et al., 2022). The formula

is defined as Equation 3:

$$y_x = \mathrm{d} + \sum_{t=1}^{p} k_t y_{x-t} + \varepsilon, \tag{3}$$

where $y_x$ is the value of the target variable at time point $t$; d is the constant term; $p$ is the order; $k_t$ is the coefficient of each autoregressive term $y_{x-t}$; and $\varepsilon$ is the error.

The AR model can predict the $PM_{2.5}$ concentrations at future moments on the basis of historical data by establishing a linear relationship between the model predicted $PM_{2.5}$ concentrations and the station observed $PM_{2.5}$ concentrations.

### 2.4 Data preprocessing and normalization

The station observed $PM_{2.5}$ concentration data were obtained from the China National Environmental Monitoring Center (http://www.cnemc.cn). Data were collected from 21 $PM_{2.5}$ monitoring stations on the northern slope of the Tianshan Mountains in 2019. The selection of data in 2019 ensures the integrity and continuity of the dataset. The hourly $PM_{2.5}$ concentration data from 00:00 to 24:00 (LST) every day for each $PM_{2.5}$ monitoring station were calculated, and the daily average value was obtained to represent the daily $PM_{2.5}$ concentration of the corresponding station. Then, the daily $PM_{2.5}$ concentration data from January to December in 2019 were obtained.

The meteorological data in 2019 used in this study including temperature, instantaneous wind speed, and pressure were obtained from the China Meteorological Data Service Centre (http://data.cma.cn/). Temperature, instantaneous wind speed, and pressure data were interpolated using the Kriging interpolation method to obtain continuous meteorological data in the study area.

In the process of model training, due to the different data ranges of variables, the feature data were normalized to stabilize the gradient changes in the network training process. The normalization formula is shown in Equation 4:

$$N_v = \frac{M - M_{\min}}{M_{\max} - M_{\min}}, \tag{4}$$

where $N_v$ is the normalized value of the variable; $M$ is the original value of the variable; $M_{\min}$ is the minimum value of the variable; and $M_{\max}$ is the maximum value of the variable.

After the feature data were normalized, the dataset was randomly split into three subsets: 60% for the training set, 20% for the validation set, and 20% for the test set. This random splitting was performed to ensure that each subset maintains a representative sample of the overall data distribution, thereby minimizing the potential bias during training and evaluation. A fixed random seed was used for this splitting process to ensure reproducibility. The validation set was specifically utilized for hyperparameter tuning, whereas the test set was reserved for final model evaluation.

### 2.5 Construction and evaluation of the GCN–TCN–AR model

#### 2.5.1 Construction of the GCN–TCN–AR model

On the basis of the above-mentioned models, we constructed the spatiotemporal framework of $PM_{2.5}$ predictions—the GCN–TCN–AR model, as shown in Figure 2. The GCN–TCN–AR model is a deep learning approach that adds an AR model to the GCN and TCN models for predicting the $PM_{2.5}$ concentrations. The GCN–TCN–AR model consists of three main parts.

In the GCN part, the model first extracted the station observed $PM_{2.5}$ concentration data and meteorological data from the spatial block into the GCN and then used the GCN model to learn the spatial association features between nodes to effectively capture the spatial dependence relationships between different monitoring stations. The output from the GCN part was then integrated with the time series data in the time block.

In the TCN part, the model used the TCN structure to learn the dynamic features of the time series data, and through the training of the TCN part, the model can better predict the future $PM_{2.5}$ concentrations. The AR modeling was performed to capture the linear characteristics of $PM_{2.5}$

concentrations at monitoring stations via the linear model. The GCN model extracted the spatial features, the TCN model extracted the complex temporal features, and the AR model extracted the linear features, which ultimately allowed the GCN–TCN–AR model to more accurately predict the $PM_{2.5}$ concentrations.

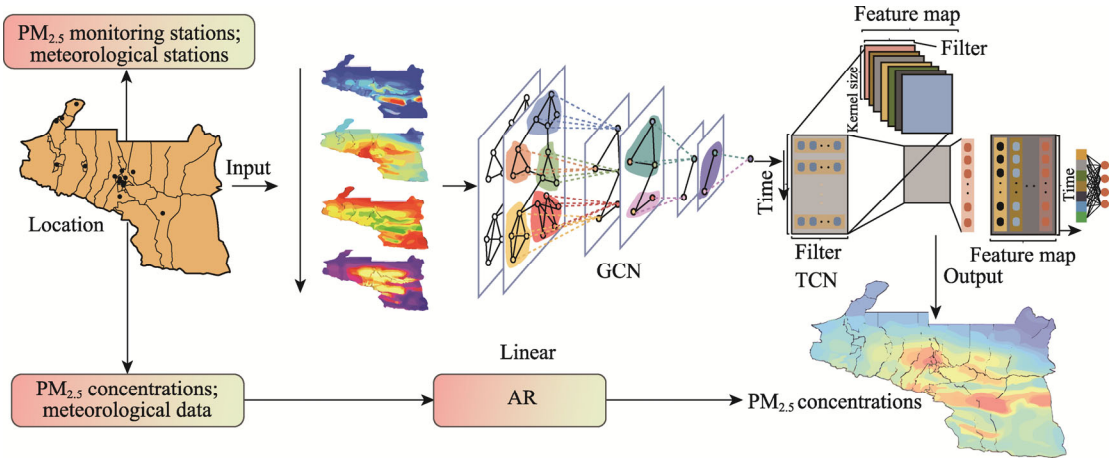


**Fig. 2** Framework using the developed GCN–TCN–AR (where GCN is the graph convolution network, TCN is the temporal convolutional network, and AR is the autoregression) model to predict $PM_{2.5}$ concentrations

### 2.5.2 Evaluation of the GCN–TCN–AR model

In this study, the following three indicators were used to evaluate the performance of the GCN–TCN–AR model. Determination of coefficient ($R^2$) is a statistic used to measure the degree of fit of the regression model. The value of $R^2$ ranges from 0.00 to 1.00, and the closer it is to 1.00, the better the model fit is. The absolute square error (MAE) and root mean square error (RMSE) are statistics used to measure the deviation between the model predicted values and the observed values; the lower the values of these indicators, the better the model performance (Liu et al., 2021). The formulas used to calculate the evaluation indicators are as follows:

$$R^2 = 1 - \frac{\sum_{o=1}^{n}(a_o - \hat{a}_o)^2}{\sum_{o=1}^{n}(a_o - \overline{a}_o)^2}, \tag{5}$$

$$\mathrm{MAE} = \frac{1}{n}\sum_{o=1}^{n}|a_o - \hat{a}_o|, \tag{6}$$

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{o=1}^{n}(a_o - \overline{a}_o)^2}, \tag{7}$$

where $n$ is the sample size; $a_o$ is the $o^{\mathrm{th}}$ observed value; $\hat{a}_o$ is the corresponding model prediction; and $\overline{a}_o$ is the mean of the observed values.

### 2.5.3 Model parameter settings

In this study, the MAE, RMSE, and accuracy rate were used as evaluation indicators to assess the performance of the GCN–TCN–AR model in $PM_{2.5}$ predictions. In the training phase of the GCN–TCN–AR model, we first adopted a large learning rate for rapid adjustment to quickly reduce losses (Qi et al., 2019). We then used a decaying learning rate to further optimize the training of the network, eventually setting the learning rate to 0.001. In addition, we set the dropout ratio to 0.1, the number of iterations to 200, and the batch size to 64. The model parameters are shown in Table 1. In the process of network training, the selection of the loss function and the optimization algorithm play important roles in predicting the $PM_{2.5}$

concentrations. To avoid overfitting, we chose the mean square error (MSE) as the loss function and used the Adam to optimize the model (Zhang et al., 2021). The Adam optimizer was selected because of its ability to dynamically adjust the learning rate, which can improve the convergence speed and stability of the model. In addition, the Adam optimizer performs well in dealing with complex lossy surfaces and can effectively address the vanishing and exploding gradients. With the above setup, we aimed to predict the $PM_{2.5}$ concentrations using the GCN–TCN–AR model with stable performance and high accuracy to improve the accuracy and reliability of air quality monitoring and prediction. To assess the model training process, we presented a plot showing the change in MSE over epochs, as well as the variation in loss during the training process (Fig. 3). The plot demonstrates a steady decrease in MSE and loss, indicating that the model effectively converges and improves its predictions over time. This suggests that the GCN–TCN–AR model is both stable and efficient in learning the underlying patterns of $PM_{2.5}$ concentrations, leading to reliable predictions.

**Table 1** Parameter settings of the GCN–TCN–AR (where GCN is the graph convolution network, TCN is the temporal convolutional network, and AR is the autoregression) model for predicting the $PM_{2.5}$ concentrations

| Parameter | Value | Parameter | Value |
|---|---|---|---|
| Number of records | 365 | Batch size | 64 |
| Number of $PM_{2.5}$ monitoring stations | 21 | Loss function | Mean square error |
| Number of meteorological stations | 15 | Learning rate | 0.001 |
| Training set | 60% | Epochs | 200 |
| Validation set | 20% | Adjusting rate | 0.6 |
| Testing set | 20% | Optimizer | Adam |

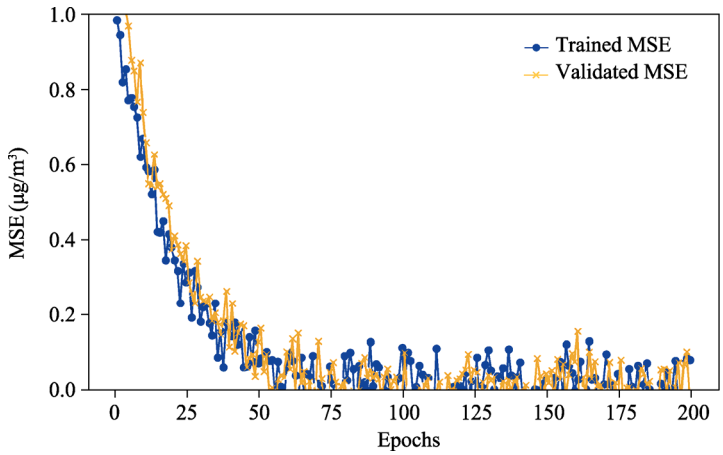


**Fig. 3** Plot showing the change in mean squared error (MSE) on the validation set with epochs and the training loss change during the training process

## 3 Results and discussion

### 3.1 Relationship between $PM_{2.5}$ concentration and meteorological factors

In addition to the influence of human activities, $PM_{2.5}$ concentration is also significantly affected by meteorological factors. The Pearson correlation coefficient between $PM_{2.5}$ concentration and each meteorological factor was subsequently calculated to obtain the degree of linear correlation between them. If the Pearson correlation coefficient is close to 1.00 or –1.00, there is a strong linear relationship between the two. If the coefficient is close to 0.00, the relationship between the two is weak, or there is no obvious linear relationship (Wang et al., 2024a).

To enhance the interpretability of the GCN–TCN–AR model, we conducted a feature

importance analysis. The results revealed that temperature and instantaneous wind speed significantly contribute to $PM_{2.5}$ predictions. These findings help us build trust in the model outputs and provide a foundation for subsequent correlation analysis. To understand the degree of relationship between $PM_{2.5}$ concentration and meteorological factors on the northern slope of the Tianshan Mountains, we used Pearson correlation to analyze the correlation heatmap of $PM_{2.5}$ concentration and meteorological factors, as shown in Figure 4. Meteorological data were derived from the daily averages of observed values from 1 January to 31 December, 2019. There were differences in the correlations between $PM_{2.5}$ concentration and meteorological factors at the seasonal scale, as shown in Table 2. Overall, meteorological factors had a very important impact on the accumulation and diffusion of $PM_{2.5}$. Instantaneous wind speed and temperature were strongly correlated with $PM_{2.5}$ concentration, whereas pressure was weakly correlated with $PM_{2.5}$ concentration (Yang et al., 2017). The $PM_{2.5}$ concentrations were significantly higher in spring and winter than in summer and autumn in the study area. Because the temperature and instantaneous wind speed were high in summer, the $PM_{2.5}$ concentrations were low in this season. In winter, both the temperature and instantaneous wind speed were low, which was not conducive to the diffusion of $PM_{2.5}$, resulting in relatively high concentrations of $PM_{2.5}$. These findings indicate that the $PM_{2.5}$ concentration is not affected by a single meteorological factor and differences in meteorological conditions may affect the $PM_{2.5}$ concentration (Chen et al., 2020).

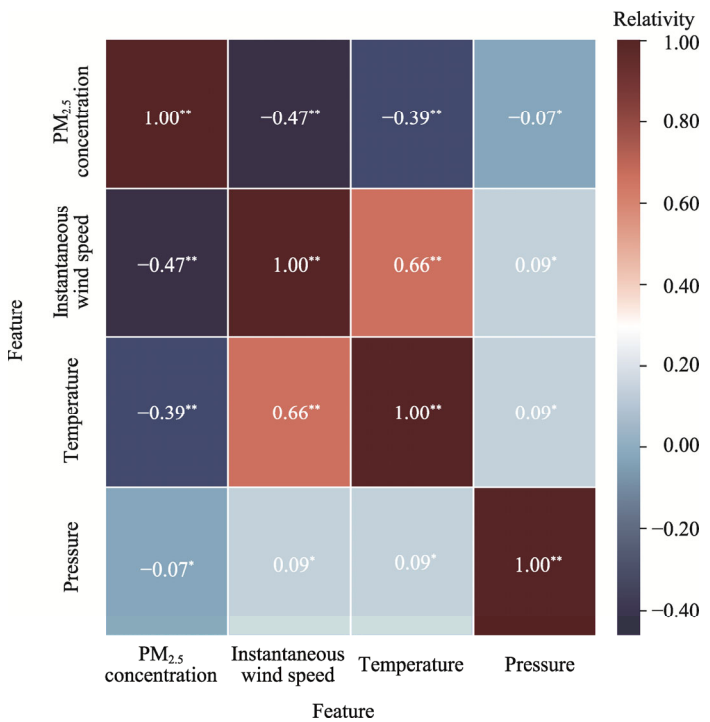


**Fig. 4** Heatmap showing the relationship between $PM_{2.5}$ concentration and meteorological factors. ** indicates a significant correlation at the $P<0.01$ level; * indicates a significant correlation at the $P<0.05$ level.

**Table 2** Correlation coefficients between $PM_{2.5}$ concentration and meteorological factors at the seasonal and annual scales

| Meteorological factor | Spring | Summer | Autumn | Winter | Annual |
|---|---|---|---|---|---|
| Instantaneous wind speed | –0.26** | –0.41** | –0.39** | 0.18** | –0.47** |
| Temperature | –0.11** | –0.34** | 0.05* | –0.05* | –0.39** |
| Pressure | –0.04 | 0.06* | –0.04 | –0.03 | –0.07* |

Note: ** indicates a significant correlation at the $P<0.01$ level; * indicates a significant correlation at the $P<0.05$ level.

### 3.2 Prediction results of $PM_{2.5}$ concentrations using the GCN–TCN–AR model

This study used the GCN–TCN–AR model to predict the $PM_{2.5}$ concentrations on the northern slope of the Tianshan Mountains from January to December in 2019. The $PM_{2.5}$ concentrations on the northern slope of the Tianshan Mountains presented a spatial distribution feature of "high in the eastern and central areas and low in the northwestern area", and the seasonal difference was also very significant, with the distribution decreasing in the order of winter>spring>autumn> summer. We forecasted the $PM_{2.5}$ concentrations at four monitoring stations in the heavily polluted areas of Urumqi, Wujiaqu, Shihezi, and Changji, as shown in Figure 5. There was a significant difference in the $PM_{2.5}$ concentrations among the different seasons at these monitoring stations. The $PM_{2.5}$ concentrations were generally higher in winter and spring, especially in Urumqi, which may be due to the accumulation of pollutants during the cold season. In contrast, the concentrations were lower in summer and autumn, possibly because of climatic factors (such as temperature and instantaneous wind speed) that help disperse and dilute air pollutants. The GCN–TCN–AR model provides relatively accurate predictions of $PM_{2.5}$ concentrations at the four monitoring stations, as evidenced by the model's predictions closely matching the actual observations (Fig. 5). Furthermore, the model effectively captures the seasonal variations in $PM_{2.5}$ concentrations, with higher concentrations in winter and spring and lower concentrations in summer and autumn, consistent with the observed patterns.

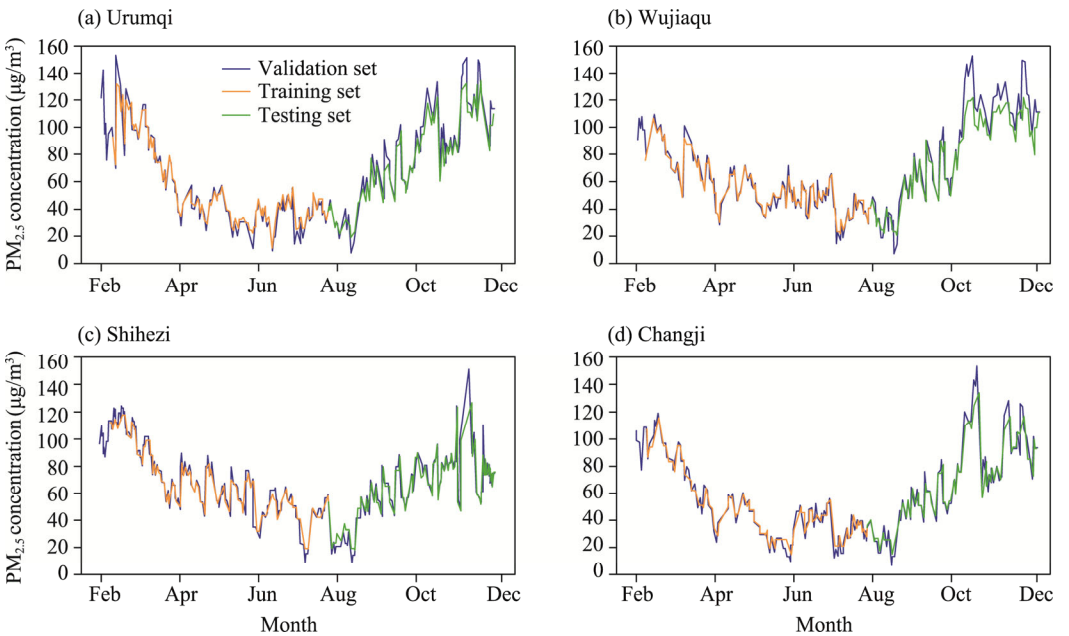


**Fig. 5** Predictions of $PM_{2.5}$ concentrations at four monitoring stations from January to December in 2019 based on the GCN–TCN–AR model. (a), Urumqi; (b), Wujiaqu; (c), Shihezi; (d), Changji.

To verify the performance of the GCN–TCN–AR model in predicting the $PM_{2.5}$ concentrations on the northern slope of the Tianshan Mountains, we drew the scatter plots of predicted $PM_{2.5}$ concentrations at Urumqi, Wujiaqu, Shihezi, and Changji monitoring stations (Fig. 6). The $R^2$ values predicted by the GCN–TCN–AR model for the four monitoring stations were 0.93, 0.91, 0.93, and 0.92, respectively, and the RMSE values were 6.85, 7.52, 7.01, and 7.28 μg/m$^3$, respectively. The trend of changes in $PM_{2.5}$ concentrations predicted by the GCN–TCN–AR model was close to that of the observed values, indicating high prediction accuracy, with the model effectively capturing the overall variation in $PM_{2.5}$ concentrations at the monitoring stations (Shi et al., 2023). This also shows that the GCN–TCN–AR model can well deal with the spatiotemporal characteristics as well as the linear and nonlinear characteristics of $PM_{2.5}$ concentrations.

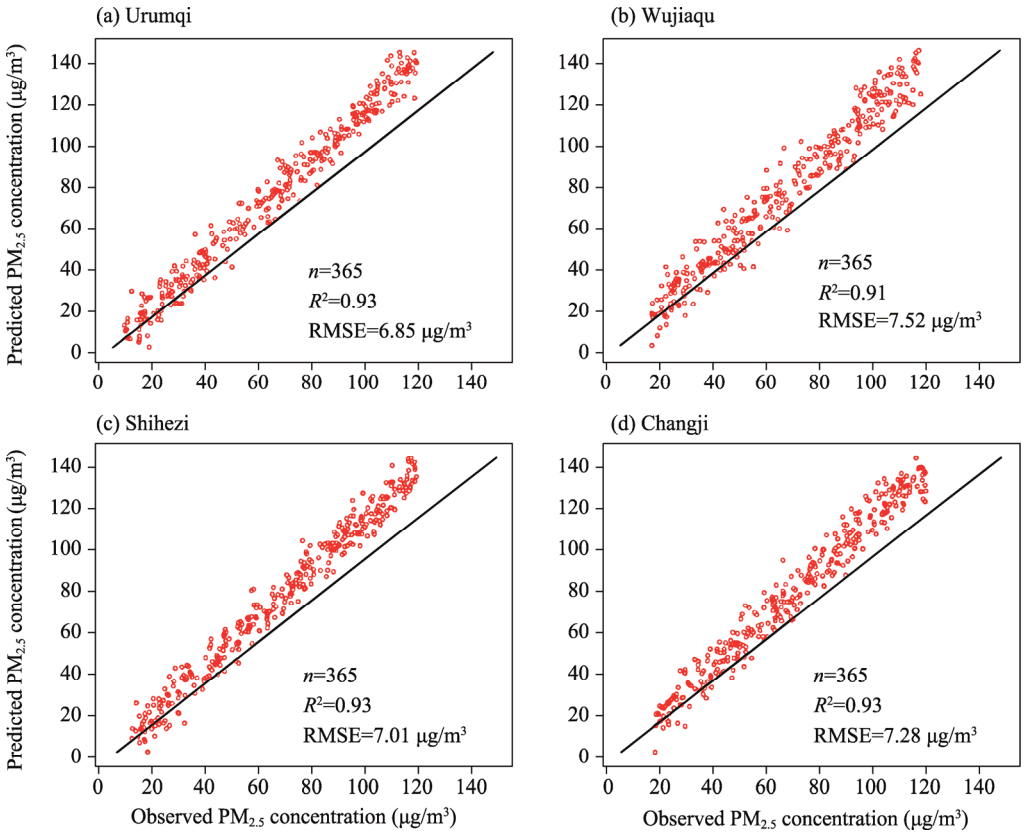


**Fig. 6** Scatter plots showing the relationship between the predicted $PM_{2.5}$ concentrations by the GCN–TCN–AR model and observed values at the daily scale in 2019 at four monitoring stations. (a), Urumqi; (b), Wujiaqu; (c), Shihezi; (d), Changji. *n*, sample size; $R^2$, determination of coefficient; RMSE, root mean square error.

### 3.3 GCN–TCN–AR model efficiency analysis

In this study, the performance of the GCN–TCN–AR model was compared with the currently popular models, including the GCN–TCN, GCN, TCN, Support Vector Regression (SVR), and AR. To ensure comparability, we used the $PM_{2.5}$ concentration dataset, which includes the observed $PM_{2.5}$ concentration data from monitoring stations, and applied the same training–testing ratio (80% for training and 20% for testing). Each model was trained under the same conditions, including hyperparameter settings and training epochs (learning rate: 0.001; batch size: 64; training epochs: 200; optimizer: Adam), to ensure a fair comparison. The RMSE was calculated on the basis of the predictions of each model on the test set. The overall performance results of each model are shown in Figure 7. The prediction effect of the AR (linear) model was clearly the worst among these neural network models, and the prediction performance of the mixed model was better than that of the single model. Compared with the GCN–TCN, GCN, TCN, SVR, and AR models, the GCN–TCN–AR model achieved the best prediction performance (Fig. 7). Specifically, the RMSE values of all the six models varied from 6.85 to 18.78 μg/m$^3$, among which the RMSE value of the GCN–TCN–AR model was the lowest (6.85 μg/m$^3$), followed by the GCN–TCN model (8.75 μg/m$^3$), while the RMSE value of the AR model was the highest (18.78 μg/m$^3$). The MAE values of all the six models ranged from 4.15 to 15.90 μg/m$^3$, among which the GCN–TCN–AR model had the lowest MAE value (4.15 μg/m$^3$), followed by the GCN–TCN model (5.97 μg/m$^3$), and the AR model had the highest value (15.90 μg/m$^3$). Figure 7 shows that the GCN–TCN–AR and GCN-TCN models that considered both spatial heterogeneity and temporal correlation, showed better performance. In contrast, the models that did not explicitly account for these factors, such as the SVR and AR models, exhibited poor performance.

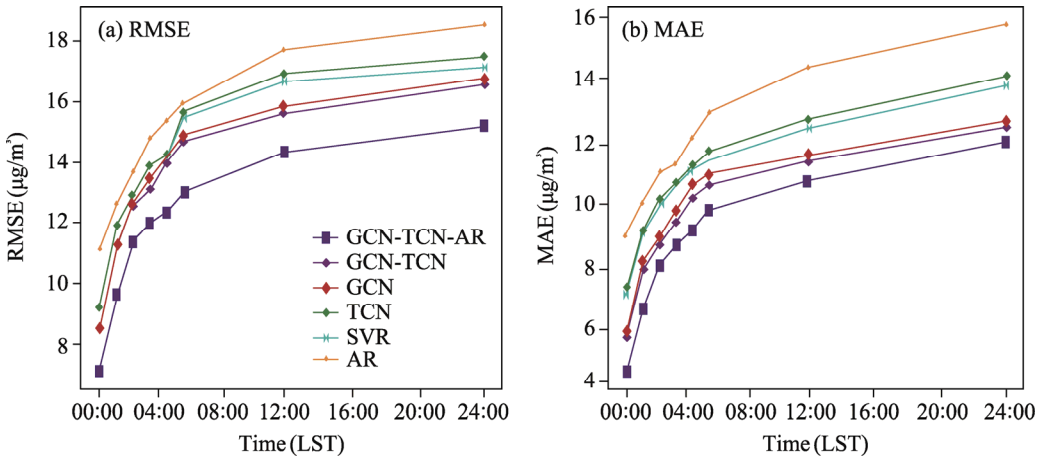


**Fig. 7** Overall performance results for models used to predict the $PM_{2.5}$ concentrations at the hourly scale. (a), RMSE; (b), MAE (mean absolute error). SVR, Support Vector Regression.

### 3.4 Gravity center analysis of $PM_{2.5}$ concentrations

In the K-means clustering analysis, the first step is to determine the number of clusters, i.e., K. The Elbow method is used to evaluate the clustering results for different K values, and the appropriate K can be selected to ensure the rationality of the clustering outcome. Based on this, the K-means algorithm was applied to cluster the $PM_{2.5}$ concentrations. The algorithm iteratively updates the cluster centers and eventually divides the samples in the region into several clusters, with each cluster representing a specific concentration feature. After clustering, the gravity center of each cluster can be calculated. The gravity center refers to the weighted average position of $PM_{2.5}$ concentrations in the region, with the specific calculation formula being the weighted average of the coordinates of each sample point and its corresponding $PM_{2.5}$ concentration. This allows for the analysis of the spatiotemporal distribution characteristics and migration trends of $PM_{2.5}$ concentrations on the northern slope of the Tianshan Mountains.

Many factors affect the $PM_{2.5}$ concentrations, such as natural factors, industrial pollution sources, population density, and automobile emissions, through which the spatial distribution characteristics of $PM_{2.5}$ concentrations can be examined (Liu et al., 2020). The K-means clustering analysis method was used to analyze the gravity center shift in the $PM_{2.5}$ concentrations on the northern slope of the Tianshan Mountains (Fig. 8). These gravity centers represented the $PM_{2.5}$ concentrations in the projections. The gravity centers were clustered into two groups, indicating that the sources of pollution were mainly located in two areas on the northern slope of the Tianshan Mountains. As shown in Figure 8, the pollution sources remained stable in 2019, with 70% of the points falling within the concentration range. The pollution sources of $PM_{2.5}$ on the northern slope of the Tianshan Mountains were mainly distributed in Urumqi City, Wujiaqu City, and Turpan City, and the interannual values of $PM_{2.5}$ concentrations were relatively high.

Urumqi and Wujiaqu cities are major industrial areas on the northern slope of the Tianshan Mountains in Xinjiang (Luo et al., 2023). Industrial development has produced large amounts of waste gas, wastewater, and solid waste, especially in the energy, fertilizer, and chemical industries. In recent years, with the increase of population density and the acceleration of urbanization, the number of cars has also increased correspondingly, the traffic congestion problem has become increasingly serious, and the emission of automobile exhaust has had an adverse impact on air quality (Lu et al., 2021). The Turpan City is located in an arid inland basin with dry climate and frequent sandstorms. This natural condition has led to high levels of sand and dust in the air, exacerbating air pollution. These factors work together to increase the severity of the environmental pollution in Turpan City (Sun et al., 2020).

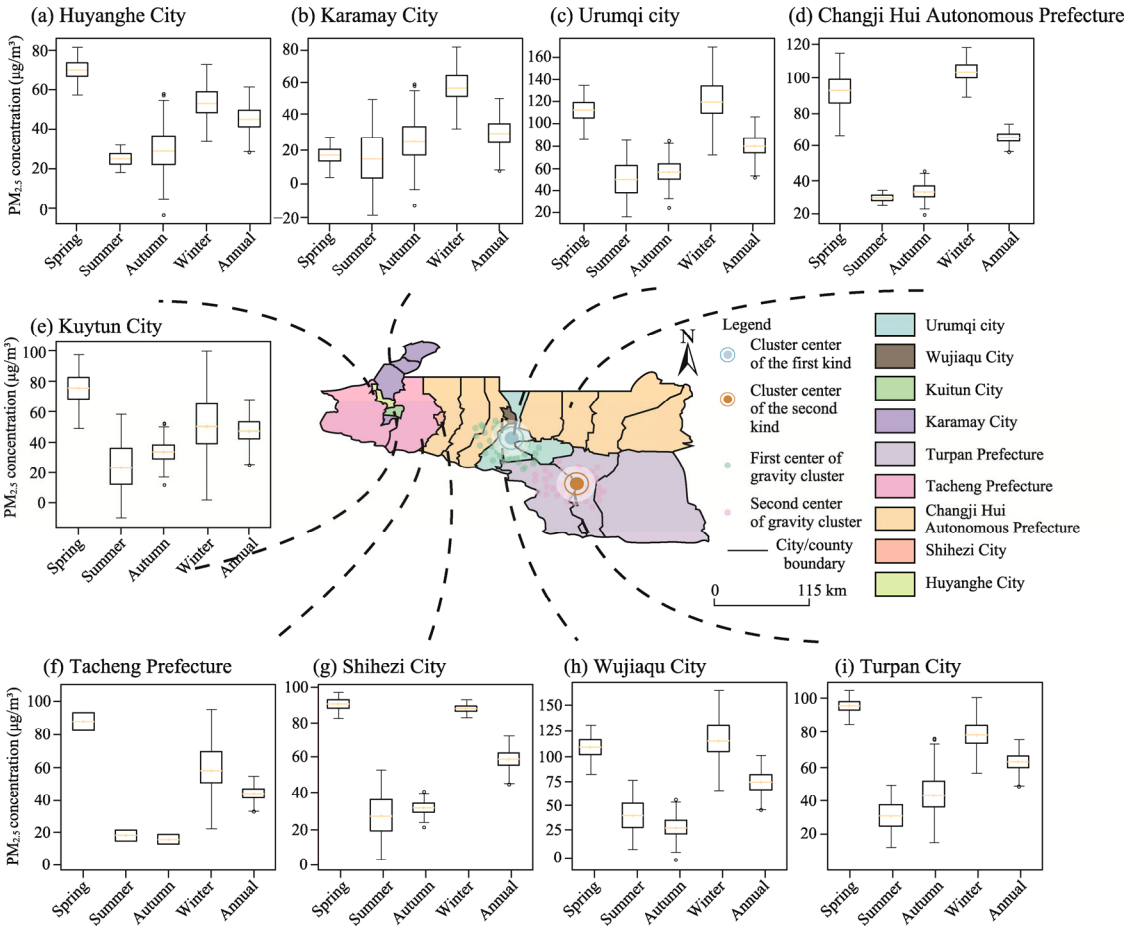


**Fig. 8** Gravity center shift in the $PM_{2.5}$ concentrations on the northern slope of the Tianshan Mountains in 2019. (a), Huyanghe City; (b), Karamay City; (c), Urumqi city; (d), Changji Hui Autonomous Prefecture; (e), Kuytun City; (f), Tacheng Prefecture; (g), Shihezi City; (h); Wujiaqu City; (i), Turpan City. The upper and lower edges of the box represent the upper quartile and the lower quartile, respectively, with the middle line representing the median; the whiskers represent the data range from the lower quartile to the minimum value and from the upper quartile to the maximum value, excluding outliers; the circles represent the outliers.

## 3.5 Distribution characteristics of $PM_{2.5}$ concentrations

### 3.5.1 Overall characteristics of $PM_{2.5}$ concentrations

$PM_{2.5}$ is an important indicator of air pollution, which is related to severe pollution in a region. The average $PM_{2.5}$ concentration on the northern slope of the Tianshan Mountains in 2019 was 50.24 μg/m$^3$ (Fig. 9). The data for Wujiaqu, Urumqi, and Shihezi cities as well as Changji Hui Autonomous Prefecture show that they had the higher pollution levels, but overall, the annual pollution level remained within the acceptable limits. The four areas with good economic development on the northern slope of the Tianshan Mountains face challenges caused by rapid economic, industrial, and urbanization processes, leading to a sharp increase in air pollutant emissions (Li et al., 2024), such as $PM_{2.5}$ emissions. Fukang and Karamay cities had low $PM_{2.5}$ concentrations, with their average annual $PM_{2.5}$ concentrations below the national standard of 35.00 μg/m³, possibly due to factors like low population density, open terrain, and high wind speed.

### 3.5.2 Annual spatial distribution characteristics of $PM_{2.5}$ concentrations

Figure 10 shows the annual spatial distribution characteristics of $PM_{2.5}$ concentrations on the northern slope of the Tianshan Mountains. In 2019, the annual average $PM_{2.5}$ concentrations on

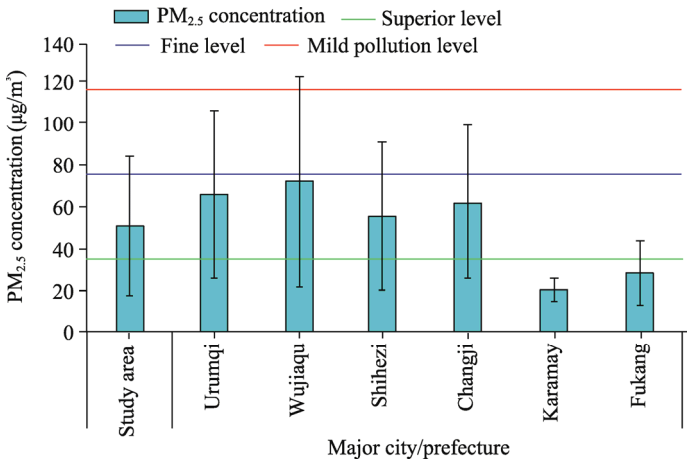


**Fig. 9** Annual average $PM_{2.5}$ concentration in the study area (the northern slope of the Tianshan Mountains) and its major cities and prefecture. Bars means standard errors.

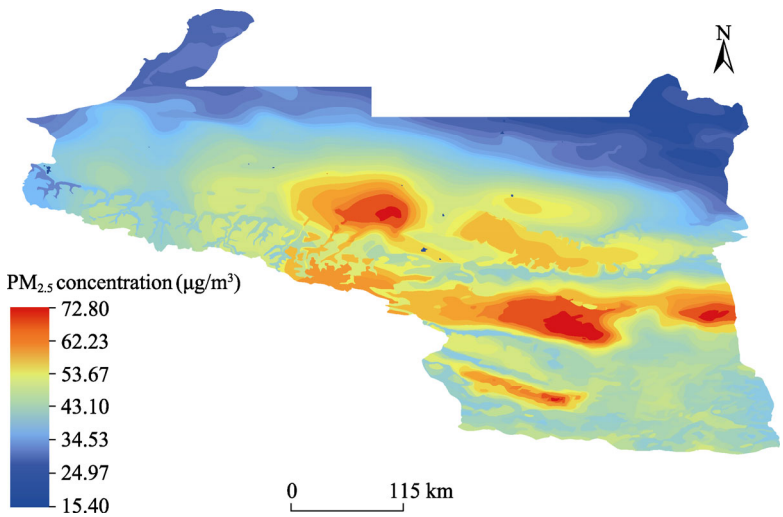


**Fig. 10** Spatial distribution characteristics of annual mean $PM_{2.5}$ concentrations on the northern slope of the Tianshan Mountains in 2019

the northern slope of the Tianshan Mountains ranged from 15.40 to 72.80 μg/m$^3$, indicating a spatial distribution characteristic of high $PM_{2.5}$ concentrations in the eastern and central areas and low $PM_{2.5}$ concentrations in the northwestern areas. According to the $PM_{2.5}$ pollution levels, we generally divided the study area into three regions, namely, the eastern region (Turpan City), the central region (Urumqi, Wujiaqu, Fukang, and Shihezi cities, as well as Changji Hui Autonomous Prefecture), and the western region (Karamay, Huyanghe, and Kuytun cities, as well as Tacheng Prefecture). In the eastern and central regions, there were two major $PM_{2.5}$ pollution centers: the eastern pollution center, with Turpan City as the core, and the central pollution center, with Urumqi and Wujiaqu cities as the cores. The annual average $PM_{2.5}$ concentrations reached 53.67–72.80 μg/m$^3$ in the eastern and central regions, extending from the two pollution centers to the surrounding areas. The annual average $PM_{2.5}$ concentrations in the western region were relatively low, at 15.40–40.15 μg/m$^3$, showing that $PM_{2.5}$ concentrations are high in the southern part and low in the northern part. Urumqi and Wujiaqu cities in the central region were the most severely polluted areas on the northern slope of the Tianshan Mountains, mainly because of their high population density, traffic exhaust emissions, and large number and scale of polluting enterprises (Peng et al., 2019). In contrast, cities in the western region had low population density

and benefited from higher instantaneous wind speed and fewer industrial and mining enterprises, resulting in low $PM_{2.5}$ concentrations (Mo et al., 2023).

### 3.5.3 Monthly spatial distribution characteristics of $PM_{2.5}$ concentrations

Figure 11 shows the monthly spatial distribution of $PM_{2.5}$ concentrations on the northern slope of the Tianshan Mountains in 2019. $PM_{2.5}$ concentrations in the study area were generally high from January to February. This pattern is consistent with that of aerosols (Guo et al., 2020). Urumqi

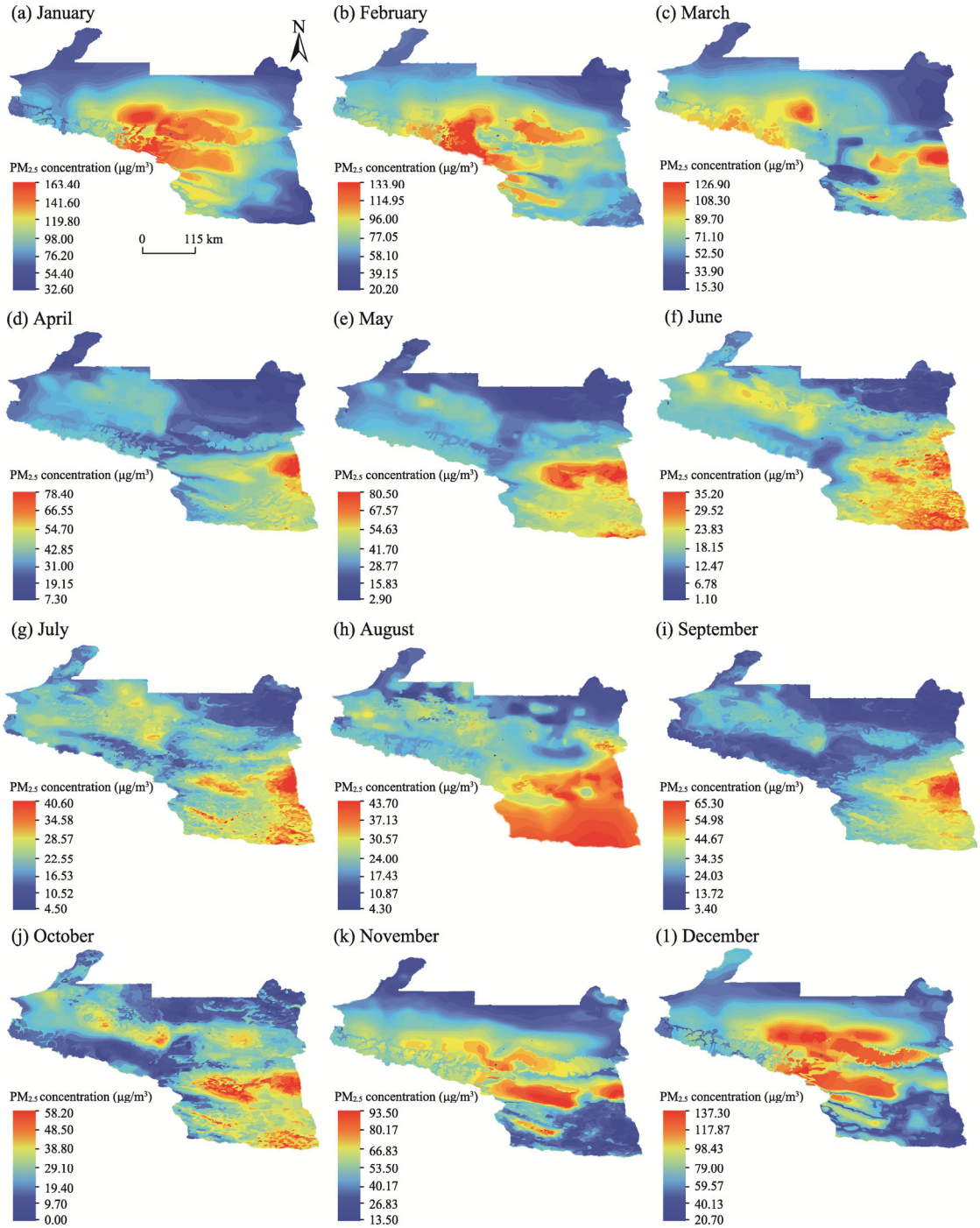


**Fig. 11** Monthly spatial distribution characteristics of $PM_{2.5}$ concentrations on the northern slope of the Tianshan Mountains from January to December in 2019 (a–l)

City, Wujiaqu City, Changji Hui Autonomous Prefecture, and northwestern Turpan City were seriously polluted, with $PM_{2.5}$ concentrations reaching 98.00–163.40 μg/m$^3$ in these two months. The finding indicates that $PM_{2.5}$ pollution was more serious on the northern slope of the Tianshan Mountains during this period, similar to the result of previous study (Ma et al., 2022). The high $PM_{2.5}$ concentrations in March were distributed mainly in Urumqi City and its surrounding western areas, including areas like Wujiaqu City and Changji Hui Autonomous Prefecture, whereas $PM_{2.5}$ concentrations in January and February spread westward. The overall spatial distribution characteristics remained basically unchanged from April to May, and $PM_{2.5}$ was concentrated in Turpan City. From June to August, $PM_{2.5}$ concentrations decreased significantly, ranging from 43.70 to 1.10 μg/m$^3$, which is related to rainfall (Yin et al., 2019). Compared with those in April and May, the distribution of $PM_{2.5}$ concentrations from June to August further expanded, revealing a spatial distribution feature of high in the southeast and low in the northwest, and the overall difference was small. In September, the spatial distribution characteristics of $PM_{2.5}$ concentrations did not change much in general but increased in some parts of Turpan City. In October, the coverage of high $PM_{2.5}$ concentrations also significantly expanded, with local increases in Urumqi City, Shihezi City, and Changji Hui Autonomous Prefecture. In November, $PM_{2.5}$ concentrations and the distribution range of high values further increased, with the peak values increasing to 13.50–93.50 μg/m$^3$; $PM_{2.5}$ concentrations in most areas on the northern slopes of the Tianshan Mountains were in the range of 53.50–93.50 μg/m$^3$, and $PM_{2.5}$ concentrations and the distribution range of low values were significantly reduced. By December, $PM_{2.5}$ concentrations and the distribution range of high values continued to increase, with high values mainly in the central region centered at Urumqi City (where the peak $PM_{2.5}$ concentration reached 137.30 μg/m$^3$) and low values (20.70–59.56 μg/m$^3$) mainly in Karamay City and southeastern Turpan City (Li et al., 2022).

## 4 Conclusions

In this work, we proposed a spatiotemporal prediction framework called the GCN–TCN–AR model that combines a GCN model, a TCN model, and an AR model. The spatiotemporal characteristics as well as linear and nonlinear features of $PM_{2.5}$ predictions were considered. An improved GCN was used to extract the spatial features of $PM_{2.5}$ concentrations. The TCN model was applied to extract the complex relationships and features of $PM_{2.5}$ time series data. An AR model was introduced to combine linear and nonlinear features to predict $PM_{2.5}$ concentrations at the monitoring stations on the northern slope of the Tianshan Mountains in 2019, which effectively improved the accuracy of regional $PM_{2.5}$ predictions.

The developed GCN–TCN–AR model predicted $R^2$ values of 0.93, 0.91, 0.93, and 0.92 at Urumqi, Wujiaqu, Shihezi, and Changji monitoring stations, respectively, as well as the RMSE values of 6.85, 7.52, 7.01, and 7.28 μg/m$^3$, respectively. This shows that the GCN–TCN–AR model can capture the spatial heterogeneity and time-varying features of $PM_{2.5}$ concentrations as well as combines linear and nonlinear features, which improves the model's accuracy and prediction ability. Compared with other neural network models (GCN–TCN, GCN, TCN, SVR, and AR), the GCN–TCN–AR model performs better.

Temperature and instantaneous wind speed were obviously negatively correlated with $PM_{2.5}$ concentrations, and pressure was not significantly negatively correlated with $PM_{2.5}$ concentrations. The overall spatial distribution characteristics of "high in the eastern and central regions and low in the northwestern region" were presented. The heavily polluted areas were concentrated in the central region of Urumqi City, Wujiaqu City, and Changji Hui Autonomous Prefecture, and the annual average $PM_{2.5}$ concentration was 50.24 μg/m$^3$ in the study area in 2019. The seasonal difference was very significant, with the distribution decreasing in the order of winter>spring>autumn>summer. However, there are still several limitations in this study. The performance of the model may limit its applicability to other regions or to the predictions of

different pollutants. Future studies could consider introducing more external factors, such as humidity and land uses, to further improve the performance of the model.

## Conflict of interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Acknowledgements

This work was supported by the Program of Support Xinjiang by Technology (2024E02028, B2-2024-0359), Xinjiang Tianchi Talent Program of 2024, the Foundation of Chinese Academy of Sciences (B2-2023-0239), and the Youth Foundation of Shandong Natural Science (ZR2023QD070). Special thanks to the anonymous reviewers and editors for helpful comments, as well as to the Xinjiang Uygur Autonomous Region Meteorological Service and Institute of Environmental Sciences of Xinjiang Uygur Autonomous Region, China for providing data.

## Author contributions

Data curation: CHEN Wenqian, BAI Xuesong, CAO Xiaoyi; Methodology: CHEN Wenqian; Software: CHEN Wenqian; Formal analysis: CHEN Wenqian; Writing - original draft: CHEN Wenqian; Visualization validation: CHEN Wenqian, BAI Xuesong; Conceptualization: CHEN Wenqian, BAI Xuesong, ZHANG Na; Writing - review & editing: CHEN Wenqian; Resources: CHEN Wenqian; Supervision: CHEN Wenqian; Project administration: CHEN Wenqian; Funding acquisition: CHEN Wenqian; Investigation: BAI Xuesong, ZHANG Na, CAO Xiaoyi. All authors approved the manuscript.

### References

Bai S N, Shen X L. 2019. $PM_{2.5}$ prediction based on LSTM recurrent neural network. Computer Application and Software, 36(1): 67–70, 104. (in Chinese)

Bhatt D, Patel C, Talsania H, et al. 2021. CNN variants for computer vision: history, architecture, application, challenges and future scope. Electronics, 10(20): 2470, doi: 10.1109/ACCESS.2021.3060744.

Chen T Q, Rubanova Y, Bettencourt J, et al. 2018. Neural ordinary differential equations. In: Bengio S. Proceedings of the 32nd International Conference on Neural Information Processing Systems (NIPS'18). NewYork: Curran Associates Inc., 6572–6583.

Chen Z Y, Chen D L, Zhao C F, et al. 2020. Influence of meteorological conditions on $PM_{2.5}$ concentrations across China: A review of methodology and mechanism. Environment International, 139: 105558, doi: 10.1016/j.envint.2020.105558.

Dong S, Wang P, Abbas K. 2021. A survey on deep learning and its applications. Computer Science Review, 40: 100379, doi: 10.1016/j.cosrev.2021.100379.

Gao X, Li W D. 2021. A graph-based LSTM model for $PM_{2.5}$ forecasting. Atmospheric Pollution Research, 12(9): 101150, doi: 10.1016/j.apr.2021.101150.

Guo B, Wang Y Q, Zhang X Y, et al. 2020. Temporal and spatial variations of haze and fog and the characteristics of $PM_{2.5}$ during heavy pollution episodes in China from 2013 to 2018. Atmospheric Pollution Research, 11(10): 1847–1856.

Jiang F X, Zhang C Y, Sun S L, et al. 2021. Forecasting hourly $PM_{2.5}$ based on deep temporal convolutional neural network and decomposition method. Applied Soft Computing, 113: 107988, doi: 10.1016/j.asoc.2021.107988.

Li K J, Talifu D, Gao B, et al. 2022. Temporal distribution and source apportionment of composition of ambient $PM_{2.5}$ in Urumqi, North-West China. Atmosphere, 13(5): 781, doi: 10.3390/atmos13050781.

Li X L, Qin D, He X L, et al. 2024. Spatial and temporal changes in land use and landscape pattern evolution in the economic belt of the northern slope of the Tianshan Mountains in China. Sustainability, 16(16): 7003, doi: 10.3390/su16167003.

Liu X P, Zou B, Feng H H, et al. 2020. Anthropogenic factors of $PM_{2.5}$ distributions in China's major urban agglomerations: A spatial-temporal analysis. Journal of Cleaner Production, 264(10): 121709, doi: 10.1016/j.jclepro.2020.121709.

Liu Y, He L J, Qin W M, et al. 2021. The effect of urban form on $PM_{2.5}$ concentration: evidence from China's 340 prefecture-level cities. Remote Sensing, 14(1): 7, doi: 10.3390/rs14010007.

Lu J, Li B, Li H, et al. 2021. Expansion of city scale, traffic modes, traffic congestion, and air pollution. Cities, 108: 102974, doi: 10.1016/j.cities.2020.102974.

Lu Y, Li K. 2023. Multistation collaborative prediction of air pollutants based on the CNN-BiLSTM model. Environmental Science and Pollution Research, 30: 92417–92435.

Luo Y T, Xu L P, Li Z Q, et al. 2023. Air pollution in heavy industrial cities along the northern slope of the Tianshan Mountains, Xinjiang: characteristics, meteorological influence, and sources. Environmental Science and Pollution Research, 30: 55092–55111.

Ma J, Ding Y X, Cheng J C, et al. 2020. Identification of high impact factors of air quality on a national scale using big data and machine learning techniques. Journal of Cleaner Production, 244: 118955, doi: 10.1016/j.jclepro.2019.118955.

Ma W, Ding J L, Wang R, et al. 2022. Drivers of $PM_{2.5}$ in the urban agglomeration on the northern slope of the Tianshan Mountains, China. Environmental Pollution, 309: 119777, doi: 10.1016/j.envpol.2022.119777.

Mo H H, You Y C, Wu L P, et al. 2023. Potential impact of industrial transfer on $PM_{2.5}$ and economic development under scenarios oriented by different objectives in Guangdong, China. Environmental Pollution, 316: 120562, doi: 10.1016/j.envpol.2022.120562.

Mohamed S A. 2019. MicroRNA detection in the pathogenesis of BAV-associated aortopathy-mediated vascular remodelling through EndMT/EMT. Journal of Internal Medicine, 285(1): 115–117.

Mohammadzadeh A K, Salah H, Jahanmahin R, et al. 2024. Spatiotemporal integration of GCN and E-LSTM networks for $PM_{2.5}$ forecasting. Machine Learning with Applications, 15: 100521, doi: 10.1016/j.mlwa.2023.100521.

Peng J B, Huang Y, Liu T, et al. 2019. Atmospheric nitrogen pollution in urban agglomeration and its impact on alpine lake-case study of Tianchi Lake. Science of the Total Environment, 688: 312–323.

Qi Y L, Li Q, Karimian H, et al. 2019. A hybrid model for spatiotemporal forecasting of $PM_{2.5}$ based on graph convolutional neural network and long short-term memory. Science of the Total Environment, 664: 1–10.

Ren Y, Wang S Y, Xia B S. 2023. Deep learning coupled model based on TCN-LSTM for particulate matter concentration prediction. Atmospheric Pollution Research, 14(4): 101703, do: 10.1016/j.apr.2023.101703.

Saha S, Gan Z T, Cheng L, et al. 2021. Hierarchical deep learning neural network (HiDeNN): an artificial intelligence (AI) framework for computational science and engineering. Computer Methods in Applied Mechanics and Engineering, 373: 113452, doi: 10.1016/j.cma.2020.113452.

Shi T, Li P Y, Yang W, et al. 2023. Application of TCN-biGRU neural network in $PM_{2.5}$ concentration prediction. Environmental Science and Pollution Research, 30: 119506–119517.

Sivarethinamohan R, Sujatha S, Priya S, et al. 2021. Impact of air pollution in health and socio-economic aspects: review on future approach. Materials Today: Proceedings, 37(2): 2725–2729.

Sun L X, Yu X, Li B S, et al. 2020. Coupling analysis of the major impact on sustainable development of the typical arid region of Turpan in Northwest China. Regional Sustainability, 1(1): 48–58.

Wang H, Gu Z J, Wang D, et al. 2024a. Evolution characteristics of Akdala $PM_{2.5}$ and correlation analysis with meteorological elements. Sichuan Environment, 43(1): 8–15. (in Chinese)

Wang J, Wu T, Mao J J, et al. 2024b. A forecasting framework on fusion of spatiotemporal features for multi-station $PM_{2.5}$. Expert Systems with Applications, 238: 121951, doi: 10.1016/j.eswa.2023.121951.

Wu X H, Song L H, Li Q L, et al. 2021. Characteristics of temporal and spatial distribution of atmospheric $PM_{2.5}$ and $PM_{10}$ in urban Taiyuan, China. Journal of Ecological Environment, 30(4): 756–762. (in Chinese)

Xia X S, Chen J J, Wang J J, et al. 2020. China $PM_{2.5}$ based on random forest model analysis of 5 factors influencing concentration. Environmental Science, 41(5): 2057–2065. (in Chinese)

Xing H T, Guo J L, Liu S A, et al. 2022. $NO_x$ emission prediction based on CNN-LSTM hybrid neural network model. Electronic Measurement, 45(2): 98–103. (in Chinese)

Yang Q Q, Yuan Q Q, Li T W, et al. 2017. The relationships between $PM_{2.5}$ and meteorological factors in China: seasonal and regional variations. International Journal of Environmental Research and Public Health, 14(12): 1510, doi: 10.3390/ijerph14121510.

Yao Y J, Gong Y G, Liu J, et al. 2023. Overview of intelligent question answering systems based on deep learning. Computer System Applications, 32(4): 1–15. (in Chinese)

Ye S, Wang P, Huang Y, et al. 2023. The spatial form of cities in the Yangtze River Delta urban agglomeration affects $PM_{2.5}$ study on the influence of spatial heterogeneity characteristics of $O_3$ pollution. Journal of Ecological Environment, 32(10): 1771–1784. (in Chinese)

Yin Z M, Cui K P, Chen S D, et al. 2019. Characterization of the air quality index for Urumqi and Turfan cities, China. Aerosol and Air Quality Research, 19(2): 282–306.

Yu Z Q, Qu Y H, Zhou G Q, et al. 2020. Numerical study on the sources of $PM_{2.5}$ pollution in the Yangtze River Delta region in autumn and winter 2018. China Environmental Science, 40(10): 4237–4246. (in Chinese)

Zeng Q L, Wang L H, Zhu S Y, et al. 2023. Long-term $PM_{2.5}$ concentrations forecasting using CEEMDAN and deep Transformer neural network. Atmospheric Pollution Research, 14(9): 101839, doi: 10.1016/j.apr.2023.101839.

Zhang M J, Wu Q Q, Zhang J, et al. 2022. Fluid micelle network for image super-resolution reconstruction. IEEE Transactions on Cybernetics, 53(1): 578–591.

Zhang S Q, Hu W, Zhao X M. 2024. Multi site air quality prediction model based on adaptive hierarchical graph convolution. Application of Computer System, 33(5): 127–135. (in Chinese)

Zhang Y W, Yuan H W, Sun X, et al. 2021. $PM_{2.5}$ concentration prediction method based on Adam attention mechanism. Journal of Atmospheric and Environmental Optics, 16(2): 117–126. (in Chinese)

Zhao G Y, He D H, Huang Y F, et al. 2021. Near-surface $PM_{2.5}$ prediction combining the complex network characterization and graph convolution neural network. Neural Computing and Applications, 33: 17081–17101.

Zhen Z, Liu J Y, Niu Y Z, et al. 2022. Analysis of factors influencing $PM_{2.5}$ in Harbin City based on multivariate time series. Journal of Henan Normal University (Natural Science Edition), 50(1): 98–107. (in Chinese)

Zhou Y L, Chang F J, Chang L C, et al. 2019. Multi-output support vector machine for regional multi-step-ahead $PM_{2.5}$ forecasting. Science of the Total Environment, 651(1): 230–240.